JOURNAL OF BALTIC SCIENCE EDUCATION

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

EFFECTIVENESS OF A GAMIFIED INTERACTIVE RESPONSE SYSTEM FOR ENHANCING PRIMARY SCIENCE EDUCATION

Yi Wen Lo

Abstract. *Primary school students of the digital generation often struggle to maintain motivation and engagement, especially under static and non-interactive teaching strategies, underscoring the need for more effective educational approaches. This quasi-experimental study examined whether integrating a gamified interactive response system, Blooket, into a sixth-grade natural science curriculum could enhance students' self-determined motivation and learning effectiveness. The experimental group (n = 52) completed online Blooket quizzes, while the control group (n = 52) took paper-and-pencil tests. Both were administered as formative assessments following the teaching sessions. Data were collected using a self-determined motivation scale and a learning effectiveness test. A self-determined motivation scale was developed specifically for this study, with strong construct validity and satisfactory model fit confirmed through exploratory and confirmatory factor analyses. The results indicated that Blooket significantly enhanced self-determined motivation and improved learning effectiveness in the experimental group. Additionally, students reported increased interest in learning and improved classroom dynamics. These findings suggest that gamified interactive response systems have the potential to enhance educational practices by effectively engaging students. Future research should examine the long-term effects and limitations of such systems across diverse educational contexts.*

Keywords: *gamified interactive response system, learning effectiveness, primary natural science education, self-determined motivation*

Yi Wen Lo
National Dong Hwa University, Taiwan

Introduction

Digital technologies profoundly influence the new generation, shaping their interactions, learning habits, and preferences. Twenge (2017) has reported that American teenagers spend an average of 2.25 hours per day texting, 2 hours browsing the internet, and 1.5 hours playing digital games, underscoring the pervasive role of technology in their daily lives. Similarly, Apaydin and Kaya (2020) have observed that students interact with digital technologies from an early age, demonstrating high levels of technological fluency. While this proficiency has provided certain advantages in educational contexts, it has also posed challenges, such as difficulty maintaining attention and a tendency to prioritize immediate gratification over long-term goals (Prensky, 2001). To address these issues, it has been suggested that educators should adapt teaching strategies to align with students' preferences for quick information uptake, visual stimuli, and interactive learning experiences (Raja & Nagasubramani, 2018).

Learning is most effective when it incorporates active engagement, clear objectives, contextual relevance, and intrinsic interest (Bruner, 1961; Bransford et al., 2000; Quinn, 2005). Cultivating intrinsic motivation is a fundamental step in fostering active learning. Self-determination theory (SDT) provides a comprehensive framework for exploring intrinsic and extrinsic motivation. According to SDT, fulfilling learners' basic psychological needs—competence, autonomy, and relatedness—enhances self-determined motivation (Deci & Ryan, 1985, 2000). When learners exhibit higher levels of self-determined motivation, their overall motivation tends to shift toward intrinsic motivation, enabling them to engage more actively in their studies. This shift not only improves the effectiveness of learning but also establishes self-determined motivation as a critical factor for achieving successful learning outcomes. Therefore, this study aimed to identify educational strategies that could enhance self-determined motivation, aiming to address the learning challenges faced by the new generation of students.

Research Focus

Plass et al. (2015) have indicated that learners often exhibit heightened motivation during gameplay, and game-based learning has been widely recognized as a form of active learning (Chen, 2019). By integrating educational content with the engaging nature of games, digital game-based learning

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

provides opportunities for students to play, experiment, make mistakes, and learn, fostering knowledge acquisition and cognitive development (Hwa, 2018). Recent studies have demonstrated that digital game-based learning environments can significantly enhance student engagement and learning outcomes across various subject areas. For example, the use of a 3D serious game to teach ancient Greek theatre has resulted in greater knowledge acquisition compared with traditional teaching methods (Papadakis et al., 2020), while integrating a game-development approach using App Inventor in computer science courses has enhanced students' programming skills and motivation through constructionist learning principles (Papadakis, 2020). However, while game-based learning and gamification share some overlapping characteristics, they represent distinct pedagogical approaches. Game-based learning involves the use of computer games explicitly designed for educational or training purposes, whereas gamification refers to the application of game elements—such as points, badges, and leaderboards—within non-game contexts to enhance motivation and learning outcomes (Liu et al., 2011; Werbach & Hunter, 2012).

Interactive Response Systems (IRSs) represent another innovative pedagogical tool. These systems enable students to provide immediate feedback via devices such as smartphones and tablets, thereby enhancing engagement, motivation, and classroom interaction (Fies & Marshall, 2006; Slain et al., 2004). Since their introduction in the 1950s, IRSs have demonstrated effectiveness in improving academic performance and pass rates (Mazur, 1997; Poulis et al., 1998). Modern online IRSs address hardware limitations of earlier models, offering enhanced functionality and reducing educators' preparation time (Green, 2014). Furthermore, incorporating game elements into Gamified Interactive Response Systems (GIRSs) helps mitigate the motivational decline associated with prolonged use of traditional IRSs (Aguilar et al., 2018).

A notable example of GIRSs is the Blooket platform, which incorporates game elements into its interactive features. Blooket's built-in game modes use mechanisms such as accruing points through correct answers, badges, and rankings to engage students in gamified learning rather than traditional gaming. Features such as immediate feedback on performance, pace regulation, and avatar recognition on leaderboards align closely with SDT principles, affirming competence, autonomy, and relatedness. These attributes make gamified tools particularly effective in fostering self-determined motivation. Moreover, the inherent structure and motivational qualities of games align well with these principles.

Despite the numerous advantages of GIRSs, prior research has predominantly focused on high school or university courses. This focus was largely due to the limitations of early systems in terms of hardware and the technological proficiency of students. In recent years, improvements in hardware availability and students' technological capabilities at the primary school level have created new opportunities for research. This study examined the impact of a GIRS on learners' self-determined motivation and learning effectiveness in the primary school natural science curriculum, offering insights into educational strategies for enhancing student engagement and attention.

Research Aim and Research Questions

This study aimed to examine the effectiveness of integrating the GIRS—Blooket into a sixth-grade natural science curriculum in enhancing students' self-determined motivation and learning effectiveness. Specifically, this study was guided by the following research questions:

1. Does the integration of Blooket into formative assessment improve students' self-determined motivation and learning effectiveness compared to paper-and-pencil tests?
2. What are the perceptions of the experimental group regarding the GIRS-integrated teaching approach?

Research Methodology

Research Design

This study employed a quasi-experimental research design using the sixth-grade natural science curriculum unit "Changes in Weather" as the teaching material. The curriculum was implemented over four weeks in November 2024, with three lessons per week, totaling 12 lessons. Both the experimental and control groups were taught by the same teacher to ensure consistency in the teaching practice. Except for the formative assessment method, all lesson materials, learning activities, lesson sequences, and teaching methods were identical for both groups.

At the end of each subunit within the "Changes in Weather" curriculum, the experimental group (two classes) completed formative assessments using the online gamified platform Blooket for approximately 15 minutes per

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

week. In contrast, the control group (two classes) took paper-and-pencil tests, each comprising 20 multiple-choice questions, as formative assessments. The test items for both groups were identical.

In the experimental group, the teacher first created a question set, selected a game mode, and set the duration for each game session. A unique game ID was then generated, which students entered to join the game. Students were able to create their own nicknames and choose avatars, thus fulfilling their psychological need for autonomy as defined by SDT. During the game, the competition results were projected in real-time onto the interactive whiteboard, enhancing the sense of competition. To win the game, students were required to answer questions correctly. Figure 1 shows a student answering a science question using the Blooket platform on a tablet device. Immediate feedback was provided after each response. Upon completion of the game, the top three rankings were displayed on the interactive whiteboard, thereby satisfying students' psychological needs for competence and relatedness, consistent with SDT. Figure 2 presents the top three performers projected onto the interactive whiteboard at the end of the Blooket game. To ensure anonymity, students' nicknames were blurred.

Figure 1
Students' Response Interface on the Blooket Platform

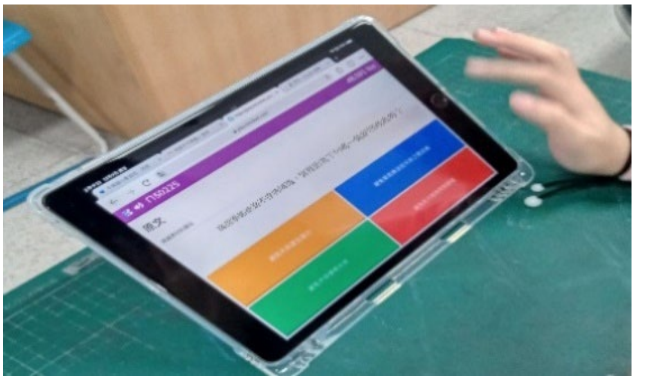
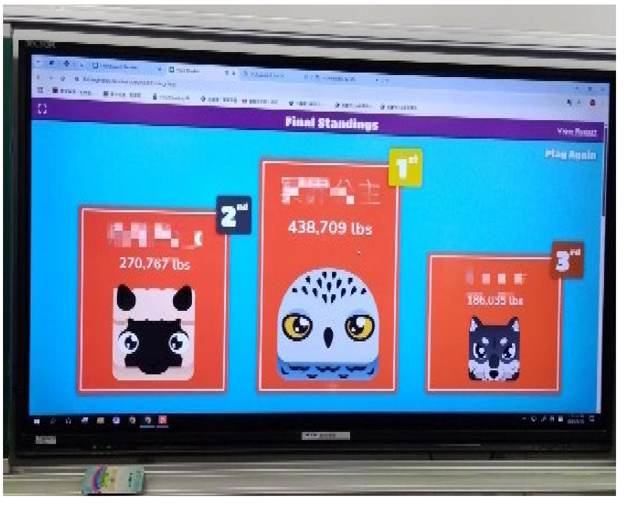


Figure 2
Top Three Performers' Avatars on the Interactive Whiteboard

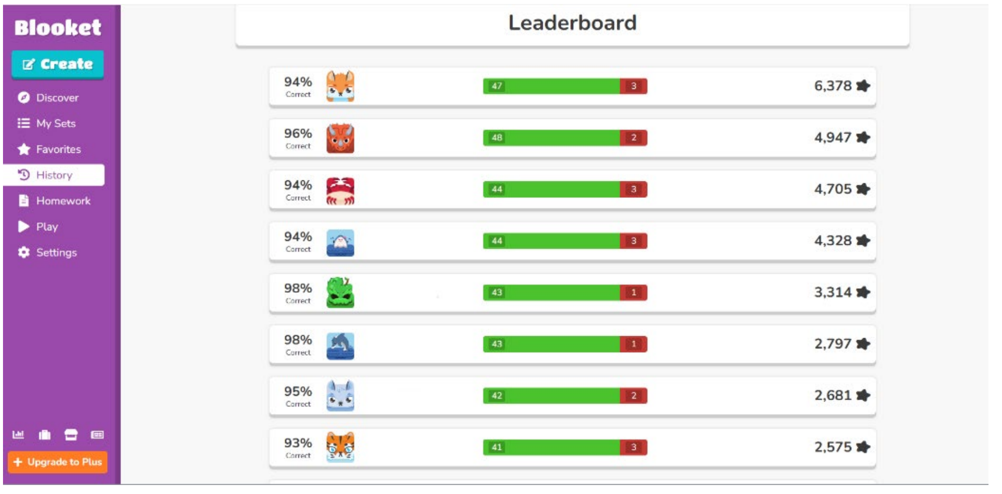


Additionally, teachers could review the platform's analytical reports after the game to assess students' response accuracy and use the information to guide teaching adjustments, as illustrated in Figure 3. The results of paper-and-pencil tests administered to the control group also served as a tool for teachers to evaluate teaching effectiveness; however, the systematic charts provided by Blooket were easier to interpret.

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Figure 3
Students' Accuracy and Performance Rankings from the Blooket Platform Report



A pretest–posttest design was adopted, wherein students in both groups completed measures of self-determined motivation and learning effectiveness before and after the intervention to examine changes attributable to variation in formative assessment methods. Learning effectiveness was measured using the Changes in Weather Test, developed specifically for this study. Self-determined motivation was assessed through the "Self-Determination Motivation Scale," validated through exploratory and confirmatory factor analyses, demonstrating strong reliability and validity.

Analysis of covariance (ANCOVA) was used to compare post-intervention differences between groups, while statistically adjusting for pretest scores. Independent samples *t tests* were conducted to assess group differences in self-determined motivation. Additionally, a Blooket Attitude Questionnaire was administered to the experimental group to assess students' perceptions of integrating Blooket into the teaching process.

Participants

This study used convenience sampling to select four sixth-grade classes from the same school. The experimental group comprised two classes with a total of 52 students (26 boys and 26 girls), while the control group consisted of two classes with 52 students (25 boys and 27 girls). All students were drawn from heterogeneously grouped classes that were formed through the school's standard random assignment process, ensuring comparable distributions in academic performance, abilities, and backgrounds. Therefore, although convenience sampling was employed, differences in students' prior experience and background across classes were minimal. Both groups were taught by the same teacher to minimize potential researcher-induced bias in the experimental implementation. All research procedures related to the experimental intervention were clearly communicated to participants and their legal guardians, including the purpose, content, and procedures of the study, as well as the voluntary nature of participation. There were no academic risks associated with participation. Anonymity was preserved, and participants' privacy and data confidentiality were strictly protected.

Additionally, one sixth-grade class with 25 students was selected for the development of the Changes in Weather Test. Convenience sampling was also used to recruit three additional primary schools to support the development of the Self-Determination Motivation Scale. Because researchers have recommended a minimum sample size of $N \geq 200$ for both exploratory factor analysis and confirmatory factor analysis to ensure stable and replicable factor structures (Comrey & Lee, 1992; de Winter et al., 2009; Kyriazos, 2018), the preliminary scale was administered to 335 fourth- and fifth-grade students at one school, resulting in 325 valid responses (response rate was 97.01%). Item analysis and exploratory factor analysis were conducted on this dataset. At the other two schools, 500 fourth- and fifth-grade students were recruited, and 458 valid responses were collected (response rate was 91.6%), on which confirmatory factor analysis was conducted. The questionnaire data collection for scale development included anonymous and low-risk items related to general learning motivation. It was conducted in accordance with approved ethical guidelines, and anonymity was strictly maintained.

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Research Instruments

Self-Determination Motivation Scale

In the past, scales developed based on SDT have been predominantly applied in areas such as sports (Ntoumanis, 2001; Mallett et al., 2007), special education (Garn et al., 2010; Lee et al., 2010), and healthcare (Williams et al., 1998). There has been limited exploration of motivational changes associated with technology integration in teaching. Therefore, a self-determination motivation scale suitable for primary school students was developed independently. This scale is grounded in the motivational categories of SDT and includes modifications to the motivational beliefs section from the Motivated Strategies for Learning Questionnaire (MSLQ), originally developed by Pintrich and De Groot (1990). The three motivational beliefs in the MSLQ—expectancy components, value components, and affective components—reflect the motivational types defined in SDT: expectancy components correspond to the psychological need for competence; value components resemble identified regulation; and affective components focus on negative emotions (test anxiety), which are linked to amotivation. For suitability to primary school students, the motivational beliefs section of the MSLQ was adapted into age-appropriate questions, and items not relevant to their everyday contexts were omitted, leading to the creation of the preliminary items for the self-determination motivation scale. The preliminary scale comprises 31 items and employs a 5-point Likert scale, ranging from "strongly agree" (5 points) to "strongly disagree" (1 point).

Item Analysis

In this study, the top 27% of the highest scores and the bottom 27% of the lowest scores were selected as extreme groups. After re-grouping, independent samples *t* tests were conducted to assess item-level discrimination. One item with a *t* value of 2.85 was removed, while the remaining items showed *t* values ranging from 4.13 to 12.78 ($p < .001$), demonstrating strong item discrimination (Shiffler, 1988).

Exploratory Factor Analysis

Kaiser (1974) proposed that before factor analysis can begin, the Kaiser-Meyer-Olkin measure of sampling adequacy (KMO) and Bartlett's test of sphericity should be used to determine whether the items are suitable for factor analysis. The KMO value of this study was .86, and the χ^2 value of Bartlett's test of sphericity was 3072.43 ($df = 325$, $p < .001$), indicating a statistically significant result. According to Kaiser (1974), a KMO value higher than .8 indicates that a scale is suitable for factor analysis.

The principal component extraction method was used to extract factors, and the maximum variation method via orthogonal rotation was used to analyze the factors. In the first factor analysis, according to the suggestion of Tabachnick and Fidell (2013), questions with factor loadings of at least .4, and preferably above .5, were considered acceptable. Thus, two questions with factor loadings lower than .4 were deleted, and one question with factor loadings spanning two factors was also deleted. For the second factor analysis of the remaining 25 questions, a total of six factors were extracted. In order to facilitate the study and analysis, these six factors were classified according to the question meaning and motivation type as follows—Self-Efficacy (SE, five items), Identified Regulation (IR, seven items), External Regulation (ER, three items), Introjected Regulation: External Expectation (IREE, three items), Introjected Regulation: Self-Demands (IRSD, three items), and Amotivation: Test Anxiety (ATA, four items). The factor loadings of each item ranged from .47 to .75, and the eigenvalues of the six factors ranged from 4.50 to 23.21, all higher than 1. The six factors could explain 60.69% of the total variation of the total scale, which was higher than the requirement of 60%, indicating that the scale had good construct validity (Hair, 2010). In the reliability analysis, the internal consistency—as indicated by Cronbach's α values of .80, .86, .70, .71, .73, and .77 for each factor, and .84 for the total scale—demonstrates good overall reliability (George & Mallery, 2003).

Confirmatory Factor Analysis

Following the analyses that shaped the formal scale, the scale was administered to a second group of 500 subjects, from which 458 valid scales were recovered. Subsequently, confirmatory factor analysis was conducted using Amos 29.0 statistical software to test the construct validity of this scale. The mean scores of the questions on the formal scale ranged from 2.98 to 4.41, with standard deviations ranging from 0.84 to 1.36. The kurtosis

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

ranged from -1.21 to 1.03, and the skewness ranged from 0.02 to -1.26, with the absolute values of skewness and kurtosis both being less than 2, indicating a normal distribution. Therefore, structural equation modeling (SEM) analysis was adopted.

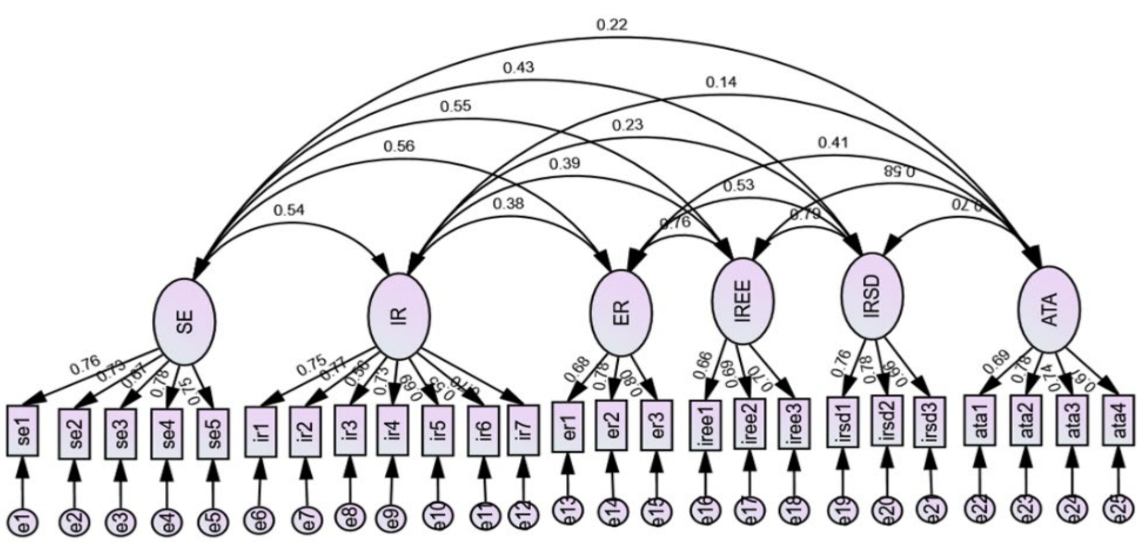
Next, the maximum likelihood method was used to perform confirmatory factor analysis on the scale items of different latent dimensions. Factor loadings ranged from .55 to .80. The overall model fit was tested using the Absolute Fit Indices, Incremental Fit Indices, and Parsimonious Fit Indices. Accordingly, the study employed the fit indices standards of Bagozzi and Yi (1988) and Hu and Bentler (1999) to conduct these tests, with the results shown in Table 1.

Table 1
Model Fit Indices for the Confirmatory Factor Analysis

Fit index (cutoff criteria)	Test value	Fit index (cutoff criteria)	Test value
χ^2 (p > .05)	544.99 (p < .05)	NFI > .90	.89
df	260	IFI > .90	.94(fit)
χ^2 / *df* < 3.00	2.10(fit)	TLI > .90	.93(fit)
RMSEA < .08	.05(fit)	CFI > .90	.94(fit)
SRMR < .08	.05(fit)	PNFI > .50	.77(fit)
GFI > .90	.92(fit)	PGFI > .50	.82(fit)
AGFI > .90	.89	CN > 200	251(fit)

From Table 1, it is evident that most indicators meet the standards. The significant χ^2 test result (p < .001) rejects the null hypothesis; however, this may be attributable to the potential inflation of the χ^2 value when the sample size is large. This was clarified using the Bollen-Stine *p*-value correction method. The AGFI indicator is sensitive to the estimated number of parameters; thus, MacCallum and Hong (1997) recommend relaxing the standard to greater than .8. Additionally, Ullman (2001) pointed out that the NFI index is sensitive to sample size and therefore suggested relaxing the standard to greater than .8. The NFI does not reflect model parsimony, which is why researchers often prefer the NNFI (TLI) index. The NNFI (TLI) value of this scale is .93, exceeding the .90 cutoff, thus meeting the fit standard. Overall, the scale's indicators show good overall model fit. Figure 4 presents the model structure and standardized parameter estimates from the confirmatory factor analysis.

Figure 4
Model Structure and Standardized Parameter Estimates from the Confirmatory Factor Analysis



https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

In terms of overall model reliability and validity, Squared Multiple Correlations (SMC) values ranged from .30 to .64, indicating a certain degree of individual reliability. Composite Reliability (CR) values ranged from .73 to .86, representing good reliability of the measurement indicators (Hair et al., 1998). Regarding validity, the Average Variance Extracted (AVE) values for the latent variables ranged from .46 to .57, indicating good convergent validity of the scale.

Overall, based on the evaluation results, the self-determination motivation scale, as extracted through exploratory factor analysis, has good model fit, CR, and convergent validity. Generally, the model's internal and external quality is good, making it suitable for studies on self-determination motivation in teaching integrated with information technology.

Changes in Weather Test

The Changes in Weather Test selected 58 multiple-choice questions from the curriculum question bank for a preliminary test with a class of 25 sixth-grade students. After the test, difficulty and discrimination indices were calculated. According to Henning's (1987) standards, questions with difficulty levels between .34 and .66 were selected. Based on Ebel's (1991) standards, questions with discrimination indices above .40 were retained, leaving 22 questions. Cronbach's α analysis was then performed, and $\alpha = .82$ indicated good internal consistency, demonstrating that the finalized test had strong reliability. The validity of the test was assessed using criterion-related validity, with the same class's fifth-grade second semester natural science final grades serving as the criterion. This yielded a Pearson correlation coefficient of $r = .64$, with $p < .001$. According to Cohen's (1988) standards, a correlation coefficient between .5 and 1 indicates a strong correlation, suggesting that the formal test effectively predicted the natural science learning effectiveness of this class.

Blooket Attitude Questionnaire

In this study, the Blooket Attitude Questionnaire was developed to assess students' perceptions, motivation, and engagement related to the use of Blooket. The questionnaire consisted of 18 close-ended items measured on a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), as well as two open-ended questions. The open-ended questions asked students to (1) describe any difficulties they encountered while using Blooket and (2) indicate their favorite game mode. The questionnaire was administered post-intervention to 52 participants in the experimental group.

Data Analysis

The data analysis employed both quantitative and qualitative approaches. Quantitative data from the pretest and posttest on self-determined motivation and learning effectiveness were analyzed using SPSS 17. Descriptive statistics (means and standard deviations) were calculated for each dimension of the Self-Determination Motivation Scale. Independent samples *t* tests were then conducted to compare the experimental and control groups, with reverse scoring applied to negatively worded motivation items (e.g., Amotivation: Test Anxiety) to ensure consistency. Additionally, ANCOVA was conducted using pretest scores as covariates to isolate the effects of the intervention.

For the Changes in Weather Test, independent samples *t* tests were also conducted to assess differences in learning effectiveness before and after the intervention. These statistical methods provided a robust quantitative evaluation of the impact of formative assessment on learning outcomes.

Qualitative data from the Blooket Attitude Questionnaire were analyzed using thematic analysis to identify key themes related to usability, enjoyment, and learning motivation. Descriptive statistics, specifically mean scores and standard deviations for the Likert scale items, were calculated to provide a quantitative summary of participants' responses.

Research Results

Self-Determination Motivation Scale

As the number of items varied across sub-scales of the Self-Determination Motivation Scale, independent samples *t* tests were conducted using the average scores of each subscale and the overall scale. Amotivation: Test

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648-3898 /Print/
ISSN 2538-7138 /Online/

Anxiety, one of the dimensions, was reverse scored. Descriptive statistics for each dimension in the pretest and posttest are presented in Tables 2 and 3.

Table 2
Descriptive Statistics for Pretest Dimensions

Dimension name	Group	*M*	*SD*	Dimension name	Group	*M*	*SD*
SE	E	2.84	0.80	IRSD	E	2.81	1.07
	C	3.21	0.79		C	3.14	1.06
IR	E	3.69	0.80	ATA	E	2.55	1.13
	C	3.84	0.84		C	3.05	1.10
ER	E	3.03	1.00	TOTAL	E	3.06	0.55
	C	3.21	0.90		C	3.33	0.45
IREE	E	2.88	1.10				
	C	3.02	1.12				

Note. E = experimental group; C = control group

From Table 2, it can be seen that the average scores for the pretest dimensions of the experimental group are generally lower than those of the control group. Further analysis using an independent samples *t* test revealed significant differences in SE ($p = .021$, $d = 0.47$), ATA ($p = .024$, $d = 0.45$), and the total scale ($p = .006$, $d = 0.54$), indicating small to moderate effects favoring the control group. This suggests that, in the pretest, the experimental group had lower self-efficacy and higher test anxiety. Overall, the self-determination motivation of the experimental group was lower than that of the control group.

As shown in Table 3, the experimental group had higher average posttest scores than the control group. The independent samples *t* tests indicated significant differences in SE ($p < .001$, $d = 1.06$), IR ($p < .001$, $d = 1.20$), ER ($p < .001$, $d = 0.96$), ATA ($p < .001$, $d = 0.91$), and the total scale ($p < .001$, $d = 1.38$), all with large effect sizes. This indicates a substantial positive impact of the intervention on students' self-determined motivation. No significant differences were found in IREE ($p = .08$, $d = 0.34$) and IRSD ($p = .37$, $d = 0.17$).

Table 3
Descriptive Statistics for Posttest Dimensions

Dimension name	Group	*M*	*SD*	Dimension name	Group	*M*	*SD*
SE	E	3.87	0.80	IRSD	E	3.26	0.79
	C	3.05	0.75		C	3.41	0.94
IR	E	4.47	0.61	ATA	E	3.75	1.26
	C	3.60	0.82		C	2.74	0.95
ER	E	4.05	1.03	TOTAL	E	3.92	0.55
	C	3.08	1.00		C	3.20	0.49
IREE	E	3.45	1.17				
	C	3.08	0.97				

Note. E = experimental group; C = control group

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Changes in Weather Test

An independent samples *t* test was conducted to analyze the pretest and posttest results of the effectiveness test. The results are shown in Table 4.

Table 4
Independent Samples t Test Results for Pretest and Posttest on Learning Effectiveness

	M	*SD*	Levene's test		Equal means t test	
			F	*p*	*t*	*p*
Pretest						
E	15.67	5.80	1.14	.289	-3.71	< .001
C	19.67	5.17				
Posttest						
E	24.92	2.69	3.92	.050	0.94	.35
C	24.33	3.72				

Note. E = experimental group; C = control group; *df* = 102; $p < .05$

As shown in Table 4, the control group ($M = 19.67$) significantly outperformed the experimental group ($M = 15.67$) in the pretest ($p < .001$, $d = 0.73$). However, no significant difference was found in posttest scores ($p = .35$, $d = 0.18$), suggesting that the experimental group caught up following the intervention. ANCOVA revealed a significant difference in adjusted posttest scores ($p = .009$, $d = 0.51$), with the experimental group ($M = 25.45$) outperforming the control group ($M = 23.81$), indicating the effectiveness of using Blooket as a teaching approach.

Experimental Group's Blooket Attitude Questionnaire

Table 5 presents descriptive statistics for the close-ended items of the Blooket Attitude Questionnaire measured on a 5-point Likert scale.

Table 5
Descriptive Statistics for the Blooket Attitude Questionnaire

Questionnaire items	*M*	*SD*
1. I look forward to playing Blooket.	4.81	0.49
2. I find Blooket fun.	4.85	0.46
3. Blooket increases my interest in the course.	4.71	0.61
4. I am focused on every question in Blooket.	4.73	0.53
5. I try to answer every question in Blooket correctly.	4.65	0.62
6. I answer each question in Blooket as quickly as possible.	4.65	0.65
7. Competing with classmates in Blooket increases my learning motivation.	4.50	0.75
8. I really want to earn rewards in Blooket.	4.56	0.75
9. Because I want to win in Blooket, I pay more attention in class.	4.37	0.84
10. I enjoy learning through Blooket.	4.63	0.66
11. Blooket helps me learn the course content and remember it longer.	4.54	0.67
12. Blooket lets me know immediately what I don't know.	4.63	0.63

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Questionnaire items	*M*	*SD*
13. I am attracted to Blooket's game interface.	4.44	0.94
14. Blooket is easy to operate.	4.54	0.78
15. I hope other subjects could also use Blooket.	4.65	0.81
16. Blooket makes the classroom atmosphere more relaxed.	4.77	0.70
17. Blooket makes me more actively participate in classroom activities.	4.60	0.72
18. Blooket enhances my enthusiasm for class discussions.	4.21	1.14

Table 5 shows that students had a positive attitude toward Blooket, with all item means exceeding 4 on a 5-point Likert scale. Students reported that Blooket increased their interest in learning, enhanced their motivation, and improved the classroom atmosphere. They also found Blooket enjoyable and easy to operate, and many expressed interest in integrating Blooket into other subjects.

In terms of qualitative responses to two open-ended questions, most students did not report any difficulties. A small number mentioned issues such as unstable internet connections (2 students), tablet crashes (1 student), misunderstanding of game rules (2 students), or playing time being too short (1 student); 46 students did not respond. These challenges were primarily related to hardware or unfamiliarity with the interface and were typically resolved through peer support.

Regarding the second open-ended question, the most frequently selected modes were Crypto Hack (10 students), Tower Defense (8 students), and Gold Quest (7 students). Students cited enjoyment, fun, competition, and game mechanics—such as stealing coins—as the main reasons for their preferences.

Discussion

Prior research has indicated that as grade levels increase, learning motivation tends to decline, a trend exacerbated by students' shorter attention spans and preference for fast-paced, visually engaging content (Eccles et al., 1993; Prensky, 2001). To address these challenges, the present study integrated Blooket—a GIRS—into a sixth-grade natural science curriculum and examined its impact on student motivation and learning effectiveness. Although the experimental group initially performed lower than the control group in both learning outcomes and self-determined motivation, post-intervention assessments showed that they often caught up with or even outperformed the control group. These findings contribute to the existing literature by demonstrating that the integration of GIRS tools such as Blooket can effectively mitigate the motivational decline observed in upper primary school students.

In the pretest phase, the experimental group exhibited significantly lower self-efficacy and higher test anxiety, with moderate effect sizes favoring the control group. However, the posttest analysis revealed substantial improvements across most motivation subscales, with statistically significant gains in SE, IR, ER, and ATA, all showing large effect sizes. This suggests a strong positive impact of Blooket-integrated teaching on students' self-determined motivation. Although no significant differences were found in IREE and IRSD, the small effect sizes suggest more limited short-term gains in these areas. Despite having clearly instructed students to respond honestly and assured them that their answers would not affect the teacher's perception, the use of self-report instruments may still be subject to social desirability bias. Future studies are encouraged to complement self-reported data with behavioral or observational measures to obtain a more comprehensive understanding of student motivation.

Regarding learning effectiveness, the control group initially outperformed the experimental group in the pretest with a large effect size. However, no significant difference emerged in the posttest (small effect size), indicating that the experimental group closed the performance gap following the intervention. Notably, ANCOVA revealed a significant difference in adjusted posttest scores favoring the experimental group, with a moderate effect size, thus supporting the effectiveness of Blooket-integrated teaching.

Student responses to the Blooket Attitude Questionnaire further supported the intervention's impact. Quantitative results showed that all mean item scores exceeded 4 on a 5-point Likert scale, reflecting generally positive perceptions. Students reported that Blooket increased their learning interest, enhanced participation motivation, and improved the classroom atmosphere. Many expressed interest in using Blooket in other subjects. Qualitative responses echoed these themes, with students frequently citing excitement, enjoyment, and the motivational nature

https://doi.org/10.33225/jbse/25.24.956

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

of competition as key reasons for their engagement. Technical issues such as internet instability or unfamiliarity with game rules were minimal and typically resolved with peer assistance.

These findings align with claims that digital-native learners are particularly responsive to interactive, game-based learning environments that provide immediate feedback and visual stimulation (Prensky, 2001). The observed motivational gains are also consistent with SDT (Deci & Ryan, 1985), which posits that environments fostering autonomy, competence, and relatedness enhance intrinsic motivation. The lack of significant differences in IREE and IRSD may reflect the deeper psychological nature of these constructs, which are less responsive to short-term interventions and may require sustained or individualized support (Koestner & Losier, 2004).

Taken together, these findings demonstrate that incorporating GIRS platforms like Blooket into primary science education can enhance learning motivation, engagement, and performance. Such tools show strong potential for addressing the evolving needs of today's learners in digitally enriched classrooms. Nevertheless, the use of convenience sampling from a single school may limit the generalizability of the findings. Future research should consider employing stratified or randomized sampling across diverse school contexts to enhance external validity.

Conclusions and Implications

From a theoretical standpoint, this study supports the applicability of SDT in technology-enhanced learning contexts. The findings underscore the importance of aligning digital teaching tools with motivational frameworks to foster more effective and engaging learning environments. Despite the observed benefits, several implementation challenges emerged. First, technical difficulties required teachers to not only possess basic troubleshooting skills but also demonstrate a willingness to adopt online IRSs. Second, the system's limited question formats—confined to multiple-choice and fill-in-the-blank—may encourage guessing and reduce cognitive engagement. Third, logistical concerns, such as device distribution and game setup, consumed valuable teaching time. These factors highlight the need for careful planning when integrating such tools into classroom practice.

To address these issues, it is recommended that schools invest in more robust network infrastructure and provide ongoing professional development in educational technology. Developers should consider adding features that reduce random guessing and enhance question diversity. Additionally, teachers may benefit from leveraging digitally skilled students to assist peers with connectivity, operation, and device management.

Future studies should examine the sustained effects of various online IRSs across subjects and grade levels to better understand their long-term educational value. Such investigations can inform the development of scalable, theory-driven digital interventions that support both motivation and learning outcomes in contemporary classrooms.

Declaration of Interest

The author declares no conflict of interest.

References

Aguilar, S. J., Holman, C., & Fishman, B. J. (2018). Game-inspired design: empirical evidence in support of gameful learning environments. *Games and Culture*, *13*(1), 44–70. https://doi.org/10.1177/1555412015600305

Apaydin, Ç., & Kaya, F. (2020). An analysis of the preschool teachers' views on alpha generation. *European Journal of Education Studies*, *6*(11), 123–141. https://www.oapub.org/edu/index.php/ejes/article/view/2815/5453

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*, 74–94. https://doi.org/10.1007/BF02723327

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. National Academy Press. https://doi.org/10.17226/9853

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, *31*, 21–32.

Chen, C. H. (2019). The impacts of peer competition-based science gameplay on conceptual knowledge, intrinsic motivation, and learning behavioral patterns. *Educational Technology Research & Development*, *67*, 179–198. https://doi.org/10.1007/s11423-018-9635-5

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Psychology Press. https://doi.org/10.4324/9781315827506

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer. https://doi.org/10.1007/978-1-4899-2271-7

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

de Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research*, *44*(2), 147–181. https://doi.org/10.1080/00273170902794206

Ebel, R. L. (1991). *Essentials of educational measurement*. Prentice Hall.

Eccles, J. S., Midgley, C., Wigfeld, A., Buchanan, C. M., Reuman, D., Flanagan, C., & MacIver, D. (1993). Development during adolescence: The impact of stage-environment fit on young adolescents' experiences in school and in families. *American Psychologist*, *48*(2), 90–101. https://doi.org/10.1037/0003-066X.48.2.90

Fies, C., & Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology*, *15*, 101–109. https://doi.org/10.1007/s10956-006-0360-1

Garn, A. C., Matthews, M. S., & Jolly, J. L. (2010). Parental influences on the academic motivation of gifted students: A self-determination theory perspective. *Gifted Child Quarterly*, *54*(4), 263–272. https://doi.org/10.1177/0016986210377657

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

Green, A. (2014). *Significant returns in engagement and performance with a free clicker app* [Working paper]. Social Science Research Network. https://doi.org/10.2139/ssrn.2447848

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Prentice Hall.

Hair, J. F. (2010). *Multivariate data analysis*. Prentice Hall.

Henning, G. (1987). *A guide to language testing: Development, evaluation, research*. Newbury House Publishers.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hwa, S. P. (2018). Pedagogical change in mathematics learning: Harnessing the power of digital game-based learning. *Educational Technology & Society*, *21*(4), 259–276.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39,* 31–36. https://doi.org/10.1007/BF02291575

Koestner, R., & Losier, G. F. (2004). Distinguishing three ways of being highly motivated: A closer look at introjection, identification, and intrinsic motivation. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 101–121). University of Rochester Press.

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Lee, Y., Wehmeyer, M. L., Palmer, S. B., Williams-Diehm, K., Davies, D. K., & Stock, S. E. (2010). The effect of student-directed transition planning with a computer-based reading support program on the self-determination of students with disabilities. *The Journal of Special Education*, *45*(2), 104-117. https://doi.org/10.1177/0022466909358916

Liu, Y., Alexandrova, T., & Nakajima, T. (2011). Gamifying intelligent environments. In *Proceedings of the 2011 International ACM Workshop on Ubiquitous Meta User Interfaces (Ubi-MUI '11)* (pp. 7–12). Association for Computing Machinery. https://doi.org/10.1145/2072652.2072655

MacCallum, R. C., & Hong, S. (1997). Power analysis in covariance structure modeling using GFI and AGFI. *Multivariate Behavioral Research*, *32*(2), 193–210. https://doi.org/10.1207/s15327906mbr3202_5

Mallett, C., Kawabata, M., Newcombe, P., Otero-Forero, A., & Jackson, S. (2007). Sport motivation scale-6 (SMS-6): A revised six-factor sport motivation scale. *Psychology of Sport and Exercise*, *8*(5), 600–614. https://doi.org/10.1016/j.psychsport.2006.12.005

Mazur, E. (1997). *Peer instruction: A user's manual*. Prentice Hall.

Ntoumanis, N. (2001). A self-determination approach to the understanding of motivation in physical education. *British Journal of Educational Psychology*, *71*(2), 225–242. https://doi.org/10.1348/000709901158497

Papadakis, S. (2020). Evaluating a game-development approach to teach introductory programming concepts in secondary education. *International Journal of Technology Enhanced Learning*, *12*(2), 127–145. https://doi.org/10.1504/ijtel.2020.10024287

Papadakis, S., Trampas, A. M., Barianos, A. K., Kalogiannakis, M., & Vidakis, N. (2020). Evaluating the learning process: The "ThimelEdu" educational game case study. In H. C. Lane, S. Zvacek, & J. Uhomoibhi (Eds.), *Proceedings of the 12th International Conference on Computer Supported Education (CSEDU 2020)* (Vol. *2*, pp. 290–298). SciTePress. https://doi.org/10.5220/0009379902900298

Pintrich, P., & DeGroot, E. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1)*,* 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Poulis, J., Massen, C., Robens, E., & Gilbert, M. (1998). Physics lecturing with audience paced feedback. *American Journal of Physics*, *66*(5), 439–441. https://doi.org/10.1119/1.18883

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon*, *9*(5), 1–6. https://doi.org/10.1108/10748120110424816

Quinn, C. N. (2005). *Engaging learning: Designing e-learning simulation games*. Pfeiffer.

Raja, R., & Nagasubramani, P. C. (2018). Impact of modern technology in education. *Journal of Applied and Advanced Research*, *3*(Suppl. 1), S33–S35. https://doi.org/10.21839/jaar.2018.v3iS1.165

Shiffler, R. E. (1988). Maximum Z scores and outliers. *The American Statistician, 42(*1), 79–80. https://doi.org/10.1080/00031305.1988.10475530

Slain, D., Abate, M., Hodges, B. M., Stamatakis, M. K., & Wolak, S. (2004). An interactive response system to promote active learning in the Doctor of Pharmacy curriculum. *American Journal of Pharmaceutical Education*, *68*(5), Article 117. https://www.hsc.wvu.edu/media/3021/interactive-response-system-slain.pdf

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

ISSN 1648–3898 /Print/
ISSN 2538–7138 /Online/

Twenge, J. M. (2017). *iGen: Why today's super-connected kids are growing up less rebellious, more tolerant, less happy—and completely unprepared for adulthood—and what that means for the rest of us.* Atria Books.
Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (4th ed., pp. 653–771). Allyn & Bacon.
Werbach, K., & Hunter, D. (2012). *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.
Williams, G. C., Freedman, Z. R., & Deci, E. L. (1998). Supporting autonomy to motivate patients with diabetes for glucose control. *Diabetes Care*, *21*(10), 1644–1651. https://doi.org/10.2337/diacare.21.10.1644

Received: *July 23, 2025* Revised: *August 26, 2025* Accepted: *September 02, 2025*

Cite as: Lo, Y. W. (2025). Effectiveness of a gamified interactive response system for enhancing primary science education. *Journal of Baltic Science Education*, *24*(5), 956–968. https://doi.org/10.33225/jbse/25.24.956

Yi Wen Lo Doctoral Student, Institute of Natural Science Education, National Dong Hwa University, No. 1, Sec. 2, Da Hsueh Rd., Shoufeng, Hualien 974301, Taiwan.
E-mail: s88b1025@gmail.com
ORCID: https://orcid.org/0000-0002-3934-0016